Rapport on Beaufort Equivalent Scales

RALF LINDAU

Institut für Meereskunde, Kiel, Germany

Abstract

The Beaufort scale derived by Lindau (1995) is recommended to be used for converting visual marine wind estimates especially for climate study purposes, where a consistent conversion of entire data sets is essential. Shortcomings of earlier Beaufort scales can be mainly explained by the statistical method of derivation, so that a major part of this report is dedicated to basic statistical considerations.

1 Introduction

Since more than one century marine meteorologists are searching for the definite conversion of Beaufort estimates into metric wind speed. In principle, the derivation procedure is rather clear. Using a suitable technique, Beaufort estimates have to be compared to reliable wind measurements in their spatial and temporal vicinity. Finding a data set of high quality marine wind measurements is, at first glance, the most crucial prerequisite for an equivalent scale. Actually, the quality of the derived scale is indeed limited by the reliability of the calibration data set. Kaufeld (1981) used wind measurements from Ocean Weather Stations (OWS) in the North Atlantic. During more than one decade three hourly (at some stations even one hourly) observations were taken continuously by professional crews. Above that, the stations were situated in the open ocean. Therefore, coastal influences on the Beaufort estimates which are intended to be calibrated can be excluded. Another advantage is that the ships stayed in general at fixed positions so that measurement errors due to the ship's speed do not occur. The huge number of observations together with the relative high accuracy qualify the wind measurements from OWS as an excellent calibration data set.

After the principal decision which data set should be used as reference, the concrete data analysis follows. How to perform this final technical step is under debate since more than hundred years. This report intends to review the discussion and to present a statistical procedure for the correct derivation of a Beaufort equivalent scale. In conclusion a concrete scale is recommended. Since questions about the appropriate statistical analysis are the most controversial part of the discussion, a detailed consideration of regression techniques is necessary.

2 Regressions

For not losing track of things, let us first consider pure linear regressions. If data pairs from two samples X and Y are available, the correlation coefficient is defined as:

$$r = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2 \sum_{i=1}^{N}(y_i - \bar{y})^2}} \tag{1}$$

which is equal to the covariance devided by the standard deviation of both samples. The regression of Y on X is defined as:

$$\hat{y} = r\,\frac{\sigma_y}{\sigma_x}(x - \bar{x}) + \bar{y} \tag{2}$$

where $\bar{x}$ and $\bar{y}$ are denoting the means of the two calibration data sets with σ_x and σ_y being their respective standard deviations. The above regression line enables us to predict individual values $\hat{y}$ for a given x; and predicting a wind speed value for a given Beaufort estimate is just what we expect from an equivalent scale.

In order to gain a better insight of the problem, it is helpful to introduce the historically used regression method, too (fig.2). For modern computers the regression line (2) is easy to calculate, but in former times it was an arduous task. Therefore, the commonly applied technique was to sort the observation pairs into classes of constant Beaufort force and to compute the mean wind speed for each of these classes. Then, the regression line of the wind speed on the Beaufort force could be obtained by connecting these class averages. For the linear case, such procedure is equivalent to the modern method. Actually, it is even more powerful since non-linear relationships are detectable, too.

As a very simple example, let us consider two thermometers T_1 and T_2 of identical type, both providing time series of the temperature at two neighbouring sites. Because of their same principal construction and their spatial proximity we suppose no bias between them and expect the same variance for both time series. Let us further assume a correlation coefficient of 0.6 between both instruments, which is caused by the small but noticable distance between each other.

As we defined a priori the universal relationship between both thermometers, a kind of equivalent scale is easy to determine here. If we should predict the measurements of T_2 from T_1, it is obvious that

$$T_2 = T_1 \tag{3}$$

would give the optimal estimate. But surprisingly, this holds true only if the characteristics of entire samples are considered. For the prediction of individual values, eq.(2) gives the best estimate. Assuming a mean temperature of 10°C, to make the example as vivid as possible, the one-sided regression of T_2 on T_1 tells us that $T_2 = 16^oC$ (fig.1) would be the best prediction for the second thermometer, if the first shows a temperature $T_1 = 20^oC$ (and $T_2 = 4^oC$, for cases when $T_1 = 0^oC$).

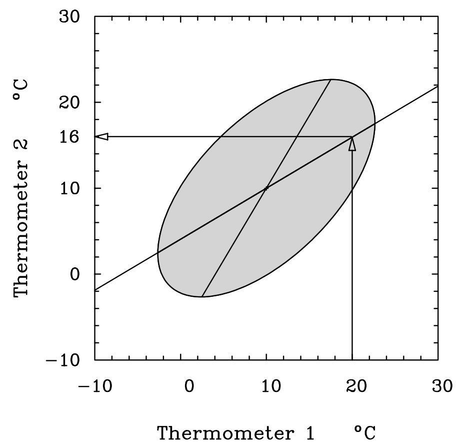

Figure 1: *Provided thermometer 1 shows 20°, the best estimate for thermometer 2 is 16°, although both instruments are neighbouring and completely identical.*

At this stage, two questions arise: As eq.(2) seems to be clearly in collision with our common sence, how can it be the optimal prediction for individual values? And if we could be convinced that this is really the case, why is eq.(2) then not the appropriate basis for an equivalent scale?

2.1 Prediction of individual values – the one-sided regressions

Let us turn to the first question. In our example, individual values can be regarded as composition of two components. Firstly, they are at least principally equal to the mean temperature of the spatially extended surrounding of both thermometers, because they can be regarded as individual realisations representative for the entire area. This is the reason why a prediction of one thermometer from the other is actually possible. Secondly, the mean temperature is modified by a stochastic spatial temperature gradient leading to slightly different values at both thermometers. Because of this variability a perfect prediction is not completely possible.

According to the above described historical method (fig.2), we can obtain the regression point by point by the following steps. Choose first a fixed value for the predictor, e.g. $T_1 = 20^oC$, sort out all temperature pairs $(T_1; T_2)$ with $T_1 = 20^oC$, and calculate the mean temperature at T_2 for these cases. As we know already from eq.(2), the result will be 16^oC.

Considering now the members of the 20^oC-class of T_1 (fig.3), we have to be aware that these values are already modified by a random deviation from their respective spatial mean. It is e.g. possible that a modified value of 20^oC results from a momentary spatial mean of 18^oC combined with a local anomaly of $+2^oC$. On the other hand, 20^oC may occur when the spatial mean for that time is 22^oC together with an anomaly of -2^oC. Since we assume the local deviations to be random, such positive and negative anomalies of the same amount have indeed the same probability. However, the point is that it are not the deviations having a different probability, but the situations itself. Extreme situa-

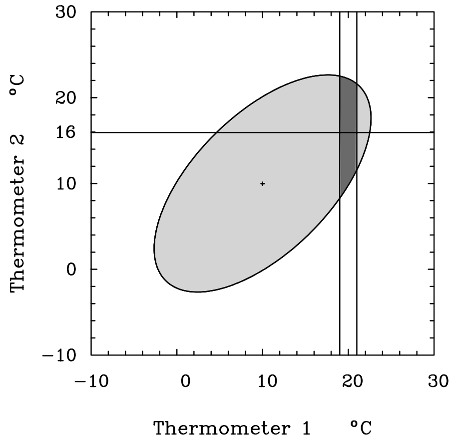

Figure 2: *The historical method to calculate regressions: Firstly, choose a value for the predictor, e.g. 20°, secondly, sort out all temperature pairs with $T_1 = 20^o$ (dark grey area), thirdly calculate the mean temperature at T_2 for these cases, finally, repeat the procedure for several predictor values and connect the results graphically.*

tions are of course less frequent than situations closer to the overall mean. Applied to our example: situations with spatial means of 18^oC are more frequent than those with 22^oC, when the overall average is 10^oC. Thus, considering the origin from which measurements of $T_1 = 20^oC$ are stemming, colder spatial means are more likely than warmer, so that 16^oC is the average of these situations.

The measurement at T_2 is just another realisation of the instantaneous temperature in the considered area. But we average over several of these values, so that T_2 reflects finally the mean temperature of the selected sample, which is 16^oC, as we have seen above, and not 20^oC. Thus, for extreme values the probability is increased that they are based solely on local events, so that they cannot be found at a neighbouring station. It is therefore wise to predict a value closer to the overall mean.

It is obvious that the example can be generalized. Substituting the expression 'spatial mean' by 'true value' and the expression 'local deviations' by 'observation errors', it will become clear that it does not matter whether real spatial differences or random observation errors are responsible for the reduced correlation coefficient.

Nevertheless, regression results similar to the above discussed are tempting sometimes to the erroneous conclusion that T_2 underestimates the temperature for warm, and overestimates it for cold situations. Obviously, this is not true, since a selection according to T_2 instead of T_1 would of course lead to the reversed result: considering only observation pairs with $T_2 = 20^oC$, it will be now T_1 which shows a mean temperature of only 16^oC.

We have seen so far that eq.(2) is indeed the best prediction for a given individual value, so that we can turn to the second question, why it should not be used as equivalent scale. I will expound in the following that such one-sided regressions do not meet the requirements of an equivalent scale, but that an improved version of eq.(3) is better suited. Both

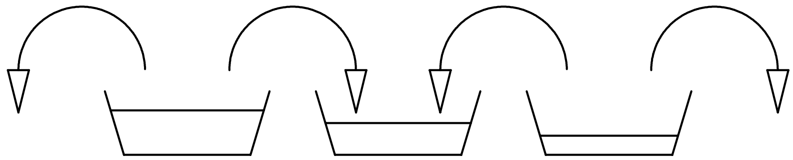

Figure 3: *Each actual measurement can be regarded as a composition of two components: The spatially mean value which is representative for a broader area, plus a random deviation for the particular site. The buckets contain situations with a mean temperature of 18, 20, and 22°, respectively. When the mean temperature is 10°C, 18°C is more frequent that 22°C. The actual temperature is a random deviation from the mean, that means in a figurative sense, splashing randomly in all directions. After this splashing procedure we examine the 20°C bucket, asking: where do these measurements come from? The probability to leave a bucket is the same for all buckets and for both directions, but the 18°C bucket is fuller so that more 'splashs' come from lower temperatures. That means if a thermometer shows 20°C, it is more likely that the surrounding is colder than 20°C.*

equations have their own advantages, and we have to face that an optimum equation for all possible applications is not attainable. A decision is necessary which of the scale characteristics are essential and which have a lower priority.

2.2 Requirements for equivalent scales – the orthogonal regression

Assuming that not individual values but an entire data set is converted by eq.(2), the disadvantages of the one-sided regression are revealed. Such theoretical data set, generated by the application of eq.(2), will contain only that part of the variance which is explained by the predictor.

The variance of the derived data set is:

$$var(\hat{y}) \;=\; \frac{1}{N-1}\sum_{i=1}^{N}(\hat{y}_i-\bar{y})^2 \tag{4}$$

From eq.(2) follows,

$$var(\hat{y}) \;=\; \frac{1}{N-1}\sum_{i=1}^{N} r^2\,\frac{\sigma_y^2}{\sigma_x^2}\,(x_i-\bar{x})^2 \tag{5}$$

which is equivalent to

$$var(\hat{y}) \;=\; r^2\,\sigma_y^2 \tag{6}$$

The loss of variance by the factor r^2 has serious consequences. It causes a substantial underestimation of the annual cycle since the correlation between wind speed and Beaufort force is perceptibly smaller than 1 (fig.4). Therefore, monthly means would be systematically underestimated for one half of the year (with anomally strong winds) and overestimated for the other half. Such performance is of course unacceptable for an equivalent

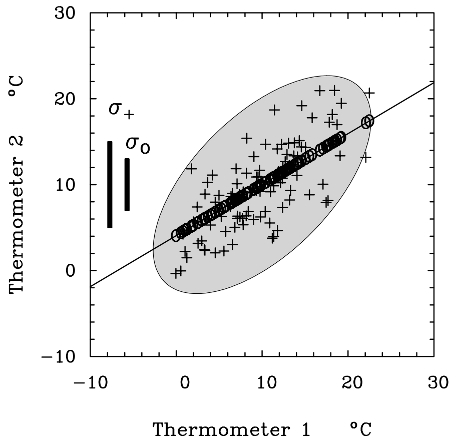

Figure 4: *Schematic figure for illustrating the reduced variance of the predicted parameter. Crosses depict real values, cycles are the prediction using the one-sided regression of measurement 2 on measurement 1. Due to the prediction, all crosses are shifted vertically, lying finally on the regression line. It is obvious that the variance is decreased by this procedure.*

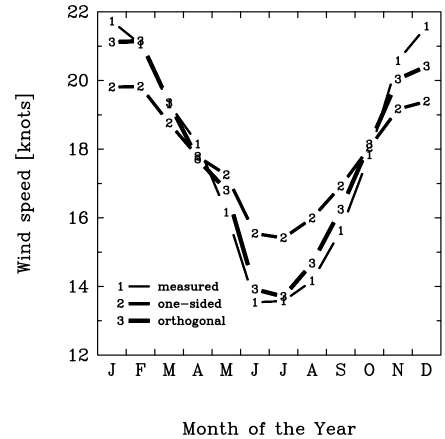

Figure 5: *The mean annual cycle of the wind speed as measured by OWS K (1). Using the one-sided regression of wind speed on Beaufort as conversion (2), the annual cycle is considerably underestimated. The orthogonal regression (3) fits much better.*

scale.

For illustration, we calcuted the two one-sided and the orthogonal regression between the wind speed measurements at OWS K and the Beaufort estimates of nearby passing merchant ships. The question is: Is it possible to predict the monthly wind speed at OWS K by the Beaufort estimates of the merchant ships by using the calculated regression lines as conversion? Figure 5 shows that the one-sided regression of wind speed on Beaufort underestimates the annual cycle seriously, while the orthogonal regression is in better agreement with the actual measurements at OWS K.

Another consequence is that one-sided regressions are necessarily not valid in other climates. Applying an equivalent scale in climate zones where it has not been derived is admittedly always a delicate venture. But using one-sided regressions, it is certain that even the longtime mean is not reproduced. If $\bar{x}$ and $\bar{\xi}$ denote the mean deriving and the mean applying Beaufort force, it follows directly from eq.(2) that the change in the obtained mean wind $\bar{\hat{\eta}} - \bar{y}$ speed will be underestimated by the factor r.

$$\frac{\bar{\hat{\eta}} - \bar{y}}{\sigma_y} = r\,\frac{\bar{\xi} - \bar{x}}{\sigma_x} \tag{7}$$

Considering again the thermometer example, another disadvantage of one-sided regressions becomes obvious. For that purpose, let us assume that one calibration attempt is carried out in winter with a mean temperature of 0^oC, and a second experiment is performed in summer with 20^oC as average. Leaving the other circumstances unchanged, the winter regression will provide $T_2 = 6^oC$ as best estimate for a given value of $T_1 = 10^oC$,

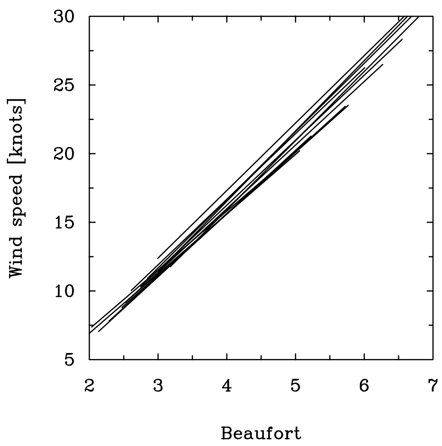

Figure 6: *The orthogonal regressions between wind measurements at OWS K and Beaufort estimates of merchant ships in the vicinity, seperately calculated for each month of the year.*

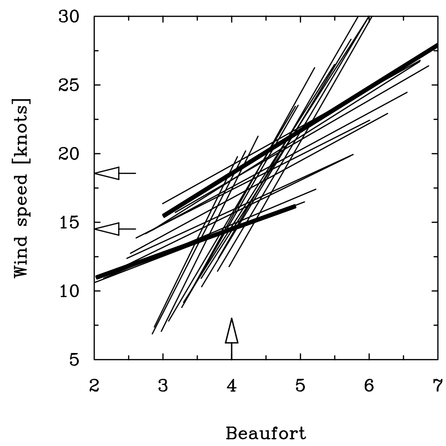

Figure 7: *As fig.6, but for the two one-sided regressions. As equivalent value for Beaufort 4, the July-regression (thick line) would give 14.5 kn, but the January-regression (thick line) 18.6 kn.*

because the correlation is assumed to be 0.6. However, the summer regression will give for the same value $(T_1 = 10^oC)$ a best estimate of $T_2 = 14^oC$. For individual predictions this is reasonable. For the wintertime, a temperature of 10^oC is a warm extreme, having opposite consequences on its probability to be representative for its surrounding as it is the case in summertime, when 10^oC is a cold extreme. Nevertheless, it is hardly acceptable that the derivation of equivalent scales leads to different results depending on the respective climate. It is not intended to deny that different wind climates might justify different equivalent scales due to changed physical conditions. But please bear in mind that absolute identical instruments were supposed in the thermometer example. Thus, obtaining two different scales is absolutely unavoidable for purely statistical reasons. Physically caused differences which are additionally possible would only modify this principle performance.

For assessing the practical consequences wind measurements at OWS K and Beaufort estimates of nearby passing merchant ships are investigated. The one-sided regression of wind speed on Beaufort, together with the reversed regression are given in fig.7. For the conversion from Beaufort force into metric wind speed, the former ones are (if one-sided regressions are used at all) appropriate. However, in summer, the equivalent value for Beaufort 4, for instance, would be 14.5 kn, considerable lower than in winter with 18.6 kn. Figure 6 shows the orthogonal regressions, seperately for each month of the year. The twelve regression lines coincide rather well, confirming that the orthogonal regressions is well suited to reflect the common relationship between wind speed and Beaufort force.

Thus, we can summarize the following. Although one-sided regressions are well suited to predict individual values, such a conversion cannot be recommended for entire data sets. Using one-sided regressions as equivalent scale, the statistical characteristics of the obtained data set will be changed substantially. The total variance will be underestimated, which causes e.g. a too weak annual cycle of the converted wind speed. For principle

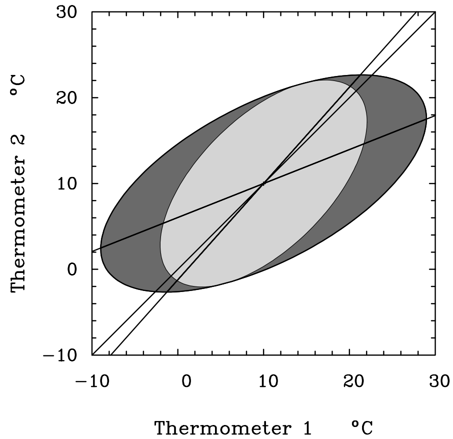

Figure 8: *Considering the thermometer example with the original data points lying in the light area, the 1–to–1 line is be the best conversion for entire data sets. This remains true, even if thermometer 1 is less accurate which would cause an elongation of the scatter ellipse (dark grey area). Computing from this data the two one-sided regressions, it becomes obvious that not the regular regression line, but the reversed one, where measurements 2 are regarded as independent, is much better suited for a conversion. The same effect would occur if thermometer 1 is not less accute, but if the variance is increased by a higher temporal resolution of the measurements.*

reasons, one-sided regressions are not applicable in other wind climates, where even the obtained total mean would be uncorrect. If different scales are derived for different climates (conceivable are twelve scales, one for each month of the year) the scales will not coincide even if Beaufort force and wind speed are actually connected by a commonly valid relationship (for which we are searching).

Hence, our first impulse using eq.(3) as conversion scale is reasonable, because we are not focussed on the optimal prediction of a single measurement, but on the conservation of the statistical characteristics. Eq.(3) is of course only valid for the above considered very simple and specific case, where equal variances and no bias were supposed. It is condensed from the following more general expression which is known as the orthogonal regression:

$$\hat{y} \;=\; \frac{\sigma_y}{\sigma_x}\,(x - \bar{x}) \;+\; \bar{y} \tag{8}$$

It is easy to show that this regression conserves actually the statistical properties discussed above. The variance of the converted data set remains unchanged, an application in other wind climates is principally possible, and calibration data sets with different total means will lead to the same results, provided that no real physical reasons are contradicting. Hence, the orthogonal regression is well suited to serve as an equivalent scale.

Nevertheless, a careful assessment of the used calibration data sets, i.e wind measurements and Beaufort estimates, is necessary. Both the temporal resolution and the relative error variance are playing here an important role (fig.8). Considering again the example of two thermometers without any systematic difference between their measurements, an equivalent scale giving the correct universal conversion should obviously have a slope of 1. This remains true even if we suppose one thermometer measuring more accurate than the other. But the unequal error variances cause different total variances for both time series so that, according to eq.(8), the slope of the orthogonal regression will not be equal to 1. A comparable effect occurs, when the standard deviations of both data sets differ due to the unequal resolutions of the considered time series. If one of both data sets contains temporally averaged values, its variance will be reduced compared to the other data set consisting of instantaneous measurements. As result, we obtain again a slope which differs from 1. In order to avoid such errors, we have to assure that the data sets used for the calibration are of the same temporal and spatial resolution so that they contain actually a comparable amount of natural variability. A second requirement is that their relative error variance has to be equal.

Hence, the orthogonal regression is the most suitable statistical way to derive an equivalent scale. But previously, the possible effects of different resolutions and different error variances of both calibration data sets have to be eliminated. Now that the principle question how to proceed is clarified, let us pass in review the numerous approaches of the last hundred years. Their assessment will show that the progress did not always take a straight course.

3 Historical Scales

In the 19th century the first attempts were made to assign metric wind speeds to the 13 wind strength classes of the british Admiral Beaufort. The principle procedure for this purpose has not changed since these days. The shipborne estimates are compared with reliable wind measurements in their temporal and spatial vicinity. The statistical analysis of these observation pairs leads then to an equivalent scale.

At the end of the 19th century knowledge about regression techniques was just evolving, but the ability of fast data processing was not available. Therefore, as mentioned in the previous chapter, the usual technique was to sort the data pairs into classes of constant Beaufort force and to compute the mean wind speed for each of these classes. Then, the one-sided regression line of the wind speed on the Beaufort force can be obtained by connecting these class averages. Reversing the sorting and the averaging parameter gives the other one-sided regression of Beaufort force on wind speed. Obviously, the second regression is suited to predict an individual Beaufort force for a given wind speed.

3.1 Köppen and Simpson

In the year 1888 a discussion began, which of the one-sided regression should be used as equivalent scale. Based on a suggestion of Köppen, Waldo (1888) proposed to take the regression of the Beaufort values on the measured wind velocity, i.e. to calculate the mean Beaufort estimate for a given wind speed class and not vice versa, which was just in contrast to the common opinion not only in these days. Even nowadays Köppen’s excelent argumentation is not always accepted, since the normally used predicting direction is from Beaufort to wind speed.

After 1888 Köppen explained in several publications his point of view (Köppen, 1916a, 1916b, 1926). In an article of the year 1916 (Köppen, 1916a) the author emphasizes that both one-sided regressions are not optimal. However, treating the measurements as independent parameter would give much better results, because the used measurements were averages over one hour, whereas the estimates were instantaneous values. Sorting the data pairs in classes of the wind speed and averaging over the Beaufort estimates would reduce the additional variance, which is included in the estimates due to the higher temporal resolution.

In London, Simpson (1906) published another Beaufort equivalent scale. Finally accepted by the WMO in 1946 as Code 1100, this scale is commonly in use until today. It is remarkable that Simpson proceeded in the same manner as suggested by Köppen. He averaged the estimates for fixed wind speed classes, thus obtaining the one-sided regression from Beaufort on wind speed. But in contrast to Köppen, Simpson considered the higher error variance of the estimates as main reason for such a data treatment.

However, both authors were aware that the variance of the Beaufort estimates is increased, may it be due to the higher temporal resolution or may it be due to the lower accuracy of the estimates, so that the regression of Beaufort on wind speed is preferable.

Already in 1916, Köppen admitted the stronger plausibility of Simpsons point of view that it were the errors which caused the higher variance of Beaufort estimates. Köppen was conviced by the fact that Curtis (1897) found no significant differences in his results when he calibrated the estimates against wind speed averages over only 10 minutes instead of hourly means.

In the beginning of the last century it was commonly accepted to use the one-sided regression of Beaufort on wind speed as equivalent scale. Köppen (1926) gave an overview and pointed out again that he was well aware of the weaknesses inherent in one-sided regressions, so that improvements were still necessary. But in those days the available data sets were too small, and may be that the experience with regression techniques were not entirely established to solve the problem definitely.

3.2 The Meteor Cruise

From 1925 to 1927, during the German Atlantic Expedition, the research vessel ’Meteor’ cruised into the South Atlantic. During this voyage the diverse problems of wind observations at sea were investigated. In this context the actual Beaufort force was hourly estimated by eight different observers, while the wind speed was recorded by anemometers at several sites of the ship. From this data set Kuhlbrodt (1936) derived a new equivalent scale. Quoting Köppen’s method of data analysis, he averaged over the Beaufort estimates, thus calculating the regression of Beaufort on wind speed. Since the ’Meteor’ touched nearly all climate zones, Kuhlbrodt computed for example a tropic and an extra-tropic scale. In order to evaluate the quality of such scales for different climates we have to keep in mind that Köppen's method is a better approximation than the reverse technique. Nevertheless, also Köppen's method leads to one-sided regression lines, which necessarily do not coincide in different climate zones even though a universal scale may exist. In the second chapter this problem is discussed in detail. Taking into account those considerations, Kuhlbrodt's attempt to derive equivalent scales for different climate zones by one-sided regressions is very questionable.

3.3 Verploegh and Richter

In the following years many other equivalent scales were derived. Verploegh (1956) used observations from two light ships at the Dutch coast. Three hourly Beaufort estimates together with anemometer measurements from 7 meters height were available for the years 1950 and 1951. After averaging the anemometer measurement over 10 minutes, Verploegh sorted the observation pairs according to the wind speed and averaged the estimates. Thus, he followed Köppen's method. However, the finally recommended scale is based not only on Verploegh's own calculations, but is an average of different scales. Among others, the results of the Meteor cruise and those of Simpson (1906) were taken into account. A scale derived by Richter (1956) was included, too. This is interesting, because Richter was one of the first who rejected Köppen's method, and returned to the antiquated procedure: to calculate the mean measured wind speed for each Beaufort class. Before merging the various scales, Verploegh discussed their differences. Richter's and his own scale showed considerable differences especially for low wind speeds, which he tried to explain purely by the actually different anemometer heights. From today's standpoint this is only half of the story. The antiquated deriving procedure of Richter leads unevitably to higher equivalent values for weak Beaufort classes, and to lower for the strong Beaufort forces. The last effect is compensated by the larger anemometer height so that only the first remains visible.

3.4 The Scientific Scale CMM-IV

In the course of the next years, the credibility of the old WMO code 1100 was more and more declining. In 1970 the Commission for Maritime Meteorology recommended a new scale, the CMM-IV, intended especially for scientific applications. Based on observations from the period 1874 to 1963, a regression line between Beaufort force and metric wind speed was calculated. But unfortunately, it was again the antiquated method which was used for derivation. Well aware that it is important which of the parameters are regarded as independent, the authors cited Köppen (1898) and Curtis (1987). But accidentially, the originally correct statement was reversed. Consequently, the low Beaufort equivalents were overestimated, while the strong ones were underestimated, and it just came true what Köppen intended to prevent by his unorthodox deriving method.

3.5 The Kaufeld Scale

Kaufeld (1981) published a new scale based on a comparison of wind speed measurements from OWS with the Beaufort estimates of nearby passing merchant ships. The large and extraordinary well suited raw data material, but even more the used regression method, gave the Kaufeld scale an outstanding importance. Kaufeld pointed out that none of the one-sided regressions is able to provide optimum results, and derived consequently a scale based principally on the linear orthogonal regression. Since in reality non-linear relationships are expected, special procedures are necessary. Kaufeld applied two concrete techniques, the construction of the angle bisection between the two one-sided regressions and the method of cumulative frequencies, both leading to similar results.

Kaufeld's proceeding is in general accordance with our recommendations to use the orthogonal regression. But for the practical application, the relative error variances of both data sets, the measurements and the estimates, have to be equal. Already Simpson and Köppen supposed that this is not case, and assessed the observation errors of Beaufort estimates to be larger than those of the measurements. Actually, this was their well-founded reason for preferring the one-sided regression of Beaufort on wind speed to the reversed regression, although they knew that both are not completely correct. Lindau (1995) showed that estimation errors are indeed larger than measurement errors, at least for measurements from OWS. After compensating these error differences, Lindau derived a new Beaufort equivalent scale. His procedure will be reviewed in the next chapter.

4 Description of the Recommended Scale

As Kaufeld, Lindau (1995) used the measurements of OWS in the North Atlantic to calibrate the Beaufort estimates from merchant ships in their vicinity. Intending to apply the orthogonal regression method, the observation errors of both data sets were calculated previously.

In order to calculate the error variance of Beaufort estimates, pairs of simultaneous ship observations are considered as a function of the distance between both ships (fig.9). With increasing distance the mean square difference between the two wind observations increases, caused by growing natural variability. As an additional component, the error variance contributes to the total variance, but it is independent from the distance, and can be regarded as a constant surcharge for each distance class. For the potential distance of zero, the natural variability vanishes and only error variance remains. As pairs of ships are considered, the double error variance appears. Repeating the analogous procedure with pairs of merchant ships and OWS, the random observation errors of the OWS measurements were concluded. It turns out that their error variance is less than half of those from merchant ship observations.

Therefore, about six instantaneous merchant ship observations in the vicinity of the OWS are averaged and compared to daily means of the OWS measurements consisting of only four individual observations. In this way, the larger errors of Beaufort estimates are exactly compensated. However, since co-located and instantaneous measurements are not available, it is unavoidable that also natural variability is included by both averaging pro-

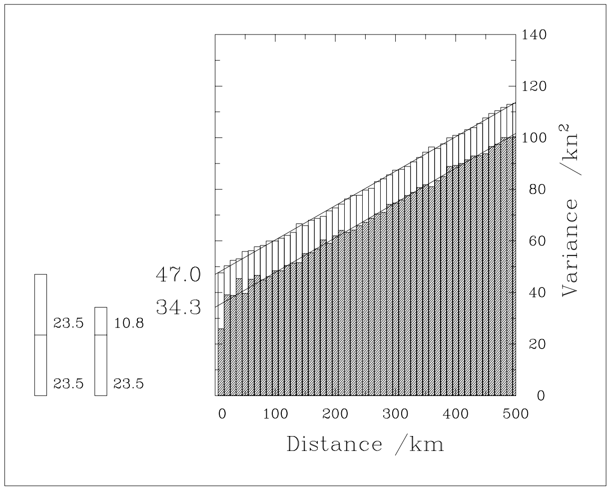

Figure 9: *Determination of the mean observation error. Mean squared wind speed differences from VOS-OWS pairs (shaded) are compared to VOS-VOS pairs (total columns) as a function of distance between the ships. For the potential distance zero, only errors contribute to the variance.*

cesses. For the OWS measurements, the temporal variability of one day is included, for the Beaufort measurements an amount of spatial variability is included which depends on the extend of the considered radius around the OWS. In order to attain comparable data sets, extactly that spatial radius is computed which corresponds to the temporal variability of one day. Depending on station and season, radii of about 300 to 400 km were found. After this procedure, Beaufort estimates and wind measurements have the same accuracy and the same resolution (in a spatial respect for the Beaufort estimates and in a temporal respect for the OWS). Applying finally the method of cumulative frequencies, the following scale was obtained. As the OWS measurements were previously reduced from 25m to 10m, the scale is valid for a height of 10m above sea level.

Bft	0	1	2	3	4	5	6	7	8	9	10	11	12
WMO	0.0	1.7	4.7	8.4	13.0	18.3	23.9	30.2	36.8	44.0	51.4	59.4	67.7
New	0.0	2.3	5.2	8.9	13.9	18.9	23.5	28.3	33.5	39.2	45.5	52.7	61.1
N	6	378	2287	8441	17197	11598	8870	4655	2068	597	122	15	1

Table 1: *New 10m–equivalent values (in knots) compared to the WMO Code 1100. N gives the number of data pairs, which consists of daily means for OWS measurements and spatial means for Voluntary Observing Ships (VOS)*

Kent & Taylor (1997) tested the performance of different Beaufort equivalent scales by comparing anemometer measured wind speeds with visual estimates, both from the Comprehesive Ocean-Atmosphere Data Set (COADS). An extraordinary meticulous height correction of the measurements was performed by using the individual anemometer heights for each ship. The agreement of converted Beaufort estimates with the corresponding measurements was checked for monthly 1°by 1°averages. In conclusion, the Beaufort scale of Lindau (1995) was found to provide the most suitable conversion for the creation

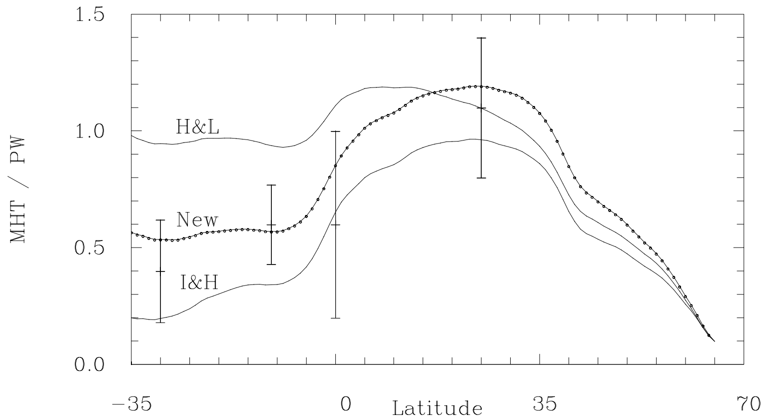

Figure 10: *Applying the Lindau scale on COADS, the turbulent fluxes at sea surface are calculated. Together with the radiative fluxes, the net energy exchange between ocean and atmosphere is concluded. The figure shows the northward meridional heat transport induced by the imbalance of the obtained net energy exchange, compared to the results of I & H: Isemer & Hasse (1987) and H & L: Hastenrath & Lamb (1978). Results of independent oceanographic sections are indicated at the relevant latitude together with their error bars: 30°S: Holfort (1994), 11°S: Speer & al. (1996), 0°: Wunsch (1984), 25°N: Bryden & Hall (1980).*

of a homogeneous monthly mean wind data set from anemometer and visual winds in COADS.

Lindau (2000) applied this scale to the marine meteorological reports of COADS. For the Atlantic Ocean, the wind dependent latent and sensible heat fluxes, together with short-wave and longwave radiation, were calculated for a 40-year period. The reliability of the resulting total net heat flux field is estimated by comparing the hereby induced meridional heat transport with independent oceanic measurements (fig.10). A good agreement was achived without any additional corrections which enhances the confidence in the used Beaufort scale.

5 Conclusions

The Beaufort scale derived by Lindau (1995) is recommended to be used for the conversion of visual estimates and metric wind speed. Especially for climate study purposes, it is essential that the characteristics of entire data sets are conserved when Beaufort estimates are converted into metric wind speed. A consistent conversion is only possible with the orthogonal regression, whereas it is the domain of one-sided regressions to give the most probable wind speed for an individual Beaufort estimate and vice versa. However, if one-sided regressions are used at all, Köppen (1898) and Simpson (1906) realized independently that the regression of Beaufort on wind speed, i.e. considering the wind speed as the independent parameter, is at least more suitable to serve as equivalent scale than the reversed regression. Larger errors in the estimates are the main reason, but different temporal resolutions of estimates and measurements, respectively, may also contribute. For a correct derivation of an equivalent scale both effects, those of different errors and those of different resolutions, must be taken into account. Lindau (1995) equalized the different errors by averaging only a small number of measurements, but a somewhat larger number of estimates. At the same time, it was ensured that the included temporal and spatial variability was equal, too. This procedure guaranteed a correct detection of the common relationship between Beaufort force and wind speed.

References

BRYDEN H.L. & M.M.HALL, 1980: Heat tranport by currents across 25°N in the Atlantic Ocean. *Science*, **207**, 884-886.

CURTIS, R.H., 1897: An attempt to determine the velocity equivalents of wind forces estimated by Beaufort's scale. *Quarterly J. Royal Meteorol. Soc.*, **23**, 24–61.

HASTENRATH & LAMB, 1978: *Heat budget atlas of the tropical Atlantic and eastern Pacific Ocean.* University of Wisconsin Press, Madison, 90pp.

HOLFORT J., 1994: Großräumige Zirkulation und meridionale Transporte im Südatlantik. *Berichte aus dem Institut für Meereskunde an der Christian-Albrechts-Universität*, Kiel, **260**, 96 pp.

ISEMER, H.J. & L.HASSE, 1987: *The Bunker climate atlas of the North Atlantic Ocean. Volume 2: Air-sea interactions.* Springer Verlag Berlin, 252 pp. KAUFELD, L., 1981: The development of a new Beaufort equivalent scale. *Meteorol. Rundschau*, **34**, 17–23.

KENT, E.C. & P.K. TAYLOR, 1997: Choice of a Beaufort Equivalent Scale *J.Atmos.Ocean.Tech*, **14**, 228–242.

KÖPPEN, W., 1898: Neuere Bestimmungen über das Verhältnis zwischen Windgeschwindigkeit und Beaufortskala. *Archiv Deutsche Seewarte*, **5**, 1–21.

KÖPPEN, W., 1916a: Beaufortskala und Windgeschwindigkeit. *Meteorol. Zeitschr.*, **33**, 88–91.

KÖPPEN, W., 1916b: Die dreizehnteilige Skala der Windstärken. *Annalen d. Hydrographie u. Marit. Met.*, **44**, 57–63.

KÖPPEN, W., 1926: Über geschätzte Windstärken und gemessene Windgeschwindigkeiten. *Annalen d. Hydrographie u. Marit. Met.*, **54**, 362–366.

KUHLBRODT, E., 1936: Vergleich geschätzter Windstärken mit gemessenen Windgeschwindigkeiten auf See. *Annalen d. Hydrographie u. Marit. Met.*, **64**, 14–23.

LINDAU, R., 1995: A new Beaufort Scale. In: Proc.Int.COADS Workshop, Kiel, Germany. *Berichte aus dem Institut für Meereskunde*, **265**, 232–252.

LINDAU, R., 2000: Climate Atlas of the Atlantic Ocean, Springer Verlag, in press.

RICHTER, J., 1956: Geschwindigkeitsäquivalente der Windstärkeschätzungen nach Beobachtungen auf deutschen Feuerschiffen. *Annalen der Meteorologie*, **7**, 267–287.

SIMPSON, C.G., 1906: The Beaufort Scale of Wind-Force. *Report of the Director of the Meteorological Office*, Offical No **180**, London

SPEER, K.G., J.HOLFORT, T.REYNAUD, G.SIEDLER, 1996: *South Atlantic Heat Transport at 11°S.* In: The South Atlantic present and past circulation, G. Welen &G. Siedler eds., Springer Verlag, Berlin, in press.

WUNSCH, C., 1984: An eclectic Atlantic Ocean circulation model. Part 1: The meridional flux of heat. *J. Phys. Ocean.*, **14,** 1712-1733.

VERPLOEGH, G., 1956: The Equivalent Velocities for the Beaufort Estimates of the Wind Force at Sea. *Koninklijk Nederlands Meteorologisch Institut, Mededelingen en Verhandelingen*, **66**, 1–38

WALDO, F., 1888: Vergleich von Beauforts Skala und Windgeschwindigkeit. *Meteorol. Zeitschr.*, **5**, 239–241.

WORLD METEOROLOGICAL ORGANISATION, 1970: The Beaufort scale of wind force (Technical and operational aspects). Commission for Marine Meteorology. *Reports to marine science affairs* Rep.No. **3**, WMO, Geneva, Switzerland, 22pp.